

(19) World Intellectual Property Organization  
International Bureau



XS

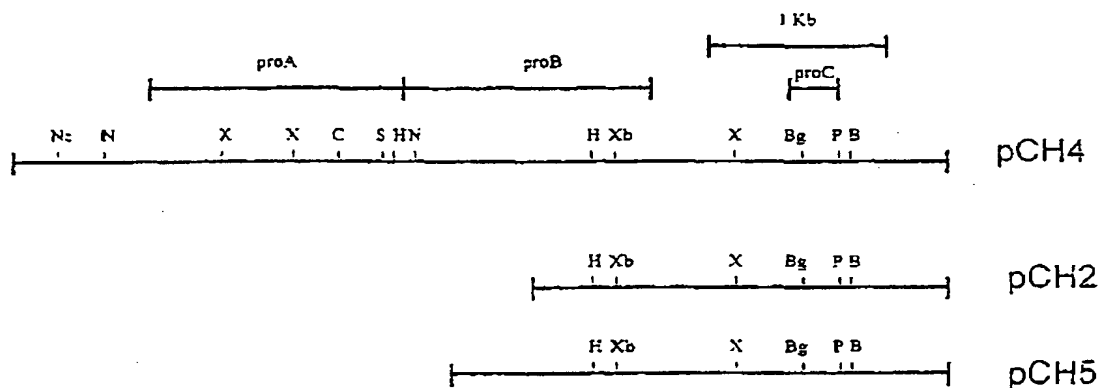
(43) International Publication Date  
1 February 2001 (01.02.2001)

PCT

(10) International Publication Number  
**WO 01/07625 A2**

- (51) International Patent Classification<sup>7</sup>: C12N 15/31, C07K 14/29, C12N 9/52, 9/02, A61K 39/00
- (21) International Application Number: PCT/US00/19763
- (22) International Filing Date: 20 July 2000 (20.07.2000)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
09/358,322 21 July 1999 (21.07.1999) US
- (71) Applicant: CORNELL RESEARCH FOUNDATION, INC. [US/US]; 20 Thornwood Drive, Suite 105, Ithaca, NY 14850 (US).
- (72) Inventor: CHIANG, Yung-fu; 204 Christopher Lane, Ithaca, NY 14850 (US).
- (74) Agents: BROWN, Michael, F. et al.; Brown, Pinnisi & Michaels, P.C., Suite 400, 118 North Tioga Street, Ithaca, NY 14850 (US).
- (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.
- (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
- Published:  
Without international search report and to be republished upon receipt of that report.
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: EHRLICHIA CANIS GENES AND VACCINES



(57) Abstract: This invention provides the sequence of 5,300 nucleotides from the E. canis genome. There are four proteins, ProA, ProB, ORF, and a cytochrome oxidase homolog, as well as a partial lipoprotein signal peptidase homolog at the carboxy terminus, coded for in this cloned fragment. The antigenic properties of these proteins allow them to be used to create a vaccine. An embodiment of this invention includes the creation of a DNA vaccine, a recombinant vaccine, and a T cell epitope vaccine.

WO 01/07625 A2

## ***Ehrlichia canis* Genes and Vaccines**

### **FIELD OF THE INVENTION**

The invention pertains to the field of veterinary pathogens. More particularly, the present invention pertains to the sequence of specific genes of the bacterial canine pathogen *Ehrlichia canis* and the application of this technology to the development of a vaccine.

### **BACKGROUND OF THE INVENTION**

The present invention relates to the sequence of genes from the *E. canis* bacterium, and the development of a vaccine against this organism.

*Ehrlichia canis* (*E. canis*) is a small gram-negative, obligately intracellular bacterium. This bacteria is the agent which causes canine monocytic ehrlichiosis (CME), a tick-borne disease which predominantly affects dogs. The most common carrier of *E. canis* is the brown dog tick *Rhipicephalus sanguineus*. The disease was described originally in Algeria in 1935. It was subsequently recognized in the United States in 1962, but is now known throughout much of the world. Canine monocytic ehrlichiosis caused much concern during the Vietnam War, when 160 military dogs died from the *E. canis* infection. There is no vaccination currently available against *E. canis*. It is a life threatening disease that continues to be an important health concern for veterinarians and pet owners alike.

Canine monocytic ehrlichiosis is an infectious blood disease. A reduction in cellular blood elements is the primary characteristic of the disease. *E. canis* lives and reproduces in the white blood cells (leukocytes). It eventually affects the entire lymphatic system, and devastates multiple organs. By targeting the white blood cells, these cells die

off rapidly. These dead blood cells migrate primarily to the spleen, which enlarges as a result. The bone marrow recognizes the loss of the white blood cells and works to form new, healthy cells. It sends out the cells prematurely, and these immature cells do not work properly. Often, these immature cells mimic those in leukemic patients, so the disease is misdiagnosed as leukemia. Canine monocytic ehrlichiosis may predispose dogs to various cancers.

There are three stages of canine monocytic ehrlichiosis. The first, acute stage mimics a mild viral infection. During the acute stage, most, if not all, of the damage is reversible and the animal is likely to recover. This is the stage where treatment is the most effective, stressing the need for early detection. Without treatment, however, the animal will progress into a subclinical (second) stage and/or to the chronic (final) stage. When the animal has reached the chronic stage, the bacterial organism has settled within the bone marrow. Many dogs in this stage suffer massive internal hemorrhage, or develop lethal complications such as sudden stroke, heart attack, renal failure, splenic rupture or liver failure.

*E. canis* can be cultured *in vitro* in a mammalian-derived cell line (DH82). Continued maintenance of these cells is difficult because the cell culture must be supplemented with primary monocytes (white blood cells found in bone marrow) every two weeks. The cultures are very slow growing, and the culture media is expensive.

Data concerning the genes in the *E. canis* genome has concentrated primarily on the 16S rRNA gene. Previous work has sequenced this gene, which is a ubiquitous component of the members of the ehrlichia family, as well as the majority of organisms worldwide. The high sequence homology between this gene throughout the living world makes it a poor candidate for vaccine development. It is necessary to find other genes within this genome if hope for a vaccine against this deadly disease can ever be realized.

Sequencing of the 16S rRNA gene indicates that *E. canis* is closely related (98.2% homology) to *E. chaffeensis*, the novel etiologic agent of human ehrlichiosis. Western blots of *E. canis* are similar when probed with antisera to *E. canis*, *E. chaffeensis* and *E.*

*ewingi* (another cause of human ehrlichiosis) indicating a close antigenic relationship between these three species (Chen *et al.*, 1994).

The indirect fluorescent antibody test (IFA) has been developed for detecting canine monocytic ehrlichiosis. IFA detects the presence of antibodies against the invading organism in a dog's blood. Unfortunately, this test is not always accurate. Sometimes, dogs will test negative in the acute phase because their immune system is delayed in forming antibodies. Another false negative may occur if there is a low titer in the chronic stage. An additional drawback of this test is the cross-reactivity found. The anti *E. canis* polyclonal antibody positively reacts with *E. chaffeensis*, undermining the specificity of the test. An alternative test, the Giesma smear, has been used to locate the actual organism in a dog's blood. Unfortunately, despite appropriate staining techniques and intensive film examination, the organisms frequently can not be located. The fallibility of these tests makes it essential to provide better diagnostic tools for this disease.

Due to difficulties in the detection of a tick bite, early diagnosis of infection, the suppression of host defenses and the nature of persistent infection of the disease, an effective vaccine against *E. canis* is urgently needed for dogs.

### SUMMARY OF THE INVENTION

This invention discloses novel sequence data for *E. canis* genes. Specifically, a clone has been identified and sequenced. Four proteins termed ProA, ProB, ORF (an open reading frame with unknown function) and a cytochrome oxidase homolog, have been identified within this clone. In addition, a partial gene encoding a lipoprotein signal peptidase homolog has been discovered.

An embodiment of this invention includes the creation of a vaccine with this sequence and protein information. The proteins disclosed in this invention are extremely antigenic. Therefore, they have the potential to be extremely useful as a vaccine. The

types of vaccine made available by this novel technology include a DNA vaccine, a recombinant vaccine, and a T cell epitope vaccine.

## BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 shows the three clones identified in the library screen.

## DESCRIPTION OF THE PREFERRED EMBODIMENT

*E. canis* causes a devastating canine disease. Currently, there is no vaccine available to prevent this disease. This invention provides the tools necessary to develop such a vaccine. More specifically, four genes have been identified from a genomic fragment of *E. canis*, named ProA, ProB, ORF and a cytochrome oxidase homolog. In addition, a partial gene coding for a lipoprotein signal peptidase homolog has been found. Any of these proteins can be utilized in an embodiment of this invention to develop a vaccine.

### Screening an *E. canis* library

To identify genes in the *E. canis* genome, a genomic DNA expression library was constructed. An *E. canis* strain isolated from dogs with canine ehrlichiosis was grown in the dog cell line DH82 by a technique being known in the art, and incorporated by reference (Dawson *et al.*, 1991; Rikihisa, 1992). The cells were harvested and the chromosomal DNA extracted as described by a technique known in the art (Chang *et al.*, 1987; Chang *et al.*, 1989a; Chang *et al.*, 1989b; Chang *et al.*, 1993a; Chang *et al.*, 1993b). To construct the library, 200 µg of DNA was partially digested with *Sau*3A. DNA fragments from 3 to 8 kb were isolated and ligated to a plasmid, pHG165 (Stewart *et al.*, 1986). The plasmids were transformed into *E. coli* TB1 (Chang *et al.*, 1987).

The library was screened with polyclonal antibodies against *E. canis*. Polyclonal antibodies were generated from dogs that had been bitten by a tick harboring *E. canis*.

The polyclonal antibodies were preabsorbed with the lysate of an *E. coli* host strain. The library was plated on petri plates at a density of 1,000 colony forming units. Colonies were transferred to nitrocellulose and each filter was probed with 1 ml of the preabsorbed polyclonal antibodies. Positive colonies were identified with a second antibody consisting of an alkaline phosphatase-conjugated goat anti-rabbit IgG (Kirkegaard and Perry Laboratories, Gaithersburg, MD), followed by color development with a substrate solution containing nitroblue tetrazolium (NBT) and 5-bromo-4-chloro-3-indolyl phosphate (BCIP). Positive clones were rescreened three times.

Three clones were isolated from this screening procedure (Figure 1). The longest genomic fragment (pCH4) encodes four complete genes and one partial gene. It completely encodes the proteins ProA, ProB, ORF and a cytochrome oxidase homolog, as well as containing the partial sequence of a lipoprotein signal peptidase homolog. ProA and ProB are located on a single operon. Restriction endonuclease digestion mapping and DNA sequencing were done by techniques known in the art, and incorporated by reference (Chang *et al.*, 1987; Chang *et al.*, 1989a; Chang *et al.*, 1989b; Chang *et al.*, 1993a; Chang *et al.*, 1993b). Briefly, the DNA sequence was determined by automated DNA sequencing on the ABI PRISM Model 377 DNA system. The complete nucleotide sequences were determined on both strands by primer walking. The thermal cycling of the sequencing reactions utilized the Taq DyeDeoxy<sup>TM</sup> Terminator Cycle sequencing kit. Databases were searched for homologous proteins through the use of the BLAST network service of the National Center for Biotechnology Information (NCBI) (Althchul *et al.*, 1990; Gish *et al.*, 1993).

#### Sequence Information

The *E. canis* genes were sequenced. The cloned fragment contains 5,300 nucleotides, and codes for four proteins. There is also one partial gene at the carboxy terminus. SEQ. ID. NO. 1 is the entire nucleotide sequence. SEQ. ID. NO. 2 and 3 are the translation of nucleotides 12 through 533 from SEQ. ID. NO. 1 and code for a cytochrome oxidase homolog. Cytochrome oxidase is important in virulence, and therefore is a strong candidate for use in a vaccine. SEQ. ID. NO. 4 and 5 are the translation of nucleotides 939 through 2,252 from SEQ. ID. NO. 1 and code for ProA. SEQ. ID. NO. 6 and 7 are the

translation of nucleotides 2,258 through 3,664 from SEQ. ID. NO. 1 and code for ProB. Preliminary evidence indicates that ProA and ProB are proteases. SEQ. ID. NO. 8 and 9 are the translation of nucleotides 4,121 through 4,795 from SEQ. ID. NO. 1 and code for ORF, a protein with unknown function. SEQ. ID. NO. 10 and 11 are the translation of the complementary sequence of nucleotides 4,884 through 5,300 from SEQ. ID. NO. 1 and code for the partial sequence of a lipoprotein signal peptidase homolog. Lipoprotein signal peptidases are membrane proteins, and by nature may be less desirable for vaccine development. However, this protein is still worth pursuing in the creation of a vaccine.

#### Overexpression of ProA, ProB, ORF, cytochrome oxidase and the lipoprotein signal peptidase homolog

The *E. canis* antigens are overexpressed in a T7 promoter plasmid. The pRSET vector allows high level expression in *E. coli* in the presence of T7 RNA polymerase, which has a strong affinity for the T7 promoter. After subcloning the antigen genes into the pRSET vector, the subclones are transformed into an F' *E. coli* JM109 strain. For maximum protein expression, the transformants are cultured to O.D. 600=0.3, exposed to IPTG (1 mM) for one hour and then transfected with M13/T7 bacteriophages at a multiplicity of infection (MOI) of 5-10 plaque forming units (pfu) per cell. Time course studies indicate that maximum induction is reached two hours after induction.

The pellet is harvested by centrifugation and the cells are resuspended in 6M Guanidinium (pH 7.8). Cells are ruptured by French press and the total lysate is spun at 6000 rpm to separate cell debris by a technique known in the art, and hereby incorporated by reference (Chang *et al.*, 1993c). Immobilized metal ion affinity chromatography (IMLAC) is used to purify each of the proteins under denaturing conditions as described by the manufacturer (Invitrogen, San Diego, CA). The protein samples are separated by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and visualized after staining with coomassie blue.

#### Vaccine Development

Prior to the present invention, no vaccine against *E. canis* had been developed. *E. canis* is endemic in dogs and closely related canidae in many parts of the world. Dogs in

North America are also increasingly at risk and the application of the present invention can potentially save the lives of thousands of dogs each year. An *E. canis* vaccine that can elicit cell-mediated immunity against this tick-borne disease of dogs is desperately needed.

### DNA Vaccine

A DNA vaccine is constructed by subcloning the gene of interest into a eukaryotic plasmid vector. Candidate vectors include, but are not limited to, pcDNA3, pCI, VR1012, and VR1020. This construct is used as a vaccine.

Each of the newly identified genes, ProA, ProB, ORF, the cytochrome oxidase homolog, or the partial lipoprotein signal peptidase homolog can be used to create a DNA vaccine (reviewed in Robinson, 1997). In addition, any immunologically active portion of these proteins is a potential candidate for the vaccine. A plasmid containing one of these genes in an expression vector is constructed. The gene must be inserted in the correct orientation in order for the genes to be expressed under the control of eukaryotic promoters. Possible promoters include, but are not limited to, the cytomegalovirus (CMV) immediate early promoter, the human tissue plasminogen activator (t-PA) gene (characterized in Degen *et al.*, 1986), and the promoter/enhancer region of the human elongation factor alpha (EF-1  $\alpha$ ) (characterized in Uetsuki *et al.*, 1989). Orientation is identified by restriction endonuclease digestion and DNA sequencing.

Expression of these gene products is confirmed by indirect immunofluorescent staining of transiently transfected COS cells. The same plasmid without these genes is used as a control. Plasmid DNA is transformed into *Escherichia coli* DH5 $\alpha$ . DNA is purified by cesium chloride gradients and the concentration is determined by a standard protocol being known in the art, and incorporated by reference (Nyika *et al.*, 1998).

Once the vector is purified, the vector containing the DNA can be suspended in phosphate buffer saline solution and directly injected into dogs. Inoculation can be done via the muscle with a needle or intravenously. Alternatively, a gene gun can be used to transport DNA-coated gold beads into cells by a technique known in the art, and hereby incorporated by reference (Fynan *et al.*, 1993). The rationale behind this type of vaccine



is that the inoculated host expresses the plasmid DNA in its cells, and produces a protein that raises an immune response. Each of the newly identified genes can be used to create a vaccine by this technique.

CpG molecules can be used as an adjuvant in the vaccine. This technique is known in the art, and is hereby incorporated by reference (Klinman *et al.*, 1997). Adjuvants are materials that help antigens or increase the immune response to an antigen. The motifs consist of an unmethylated CpG dinucleotide flanked by two 5' purines and two 3' pyrimidines. Oligonucleotides containing CpG motifs have been shown to activate the immune system, thereby boosting an antigen-specific immune response. This effect can be utilized in this invention by mixing the CpG oligonucleotides with the DNA vaccine, or physically linking the CpG motifs to the plasmid DNA.

#### Recombinant Vaccine

In order to develop a recombinant vaccine, each of the genes is individually subcloned into overexpression vectors, and then purified for vaccine development. ProA, ProB, ORF, the cytochrome oxidase homolog or the partial lipoprotein signal peptidase homolog is expressed in a plasmid with a strong promoter such as the tac, T5, or T7 promoter. Alternatively, immunologically active fragments of these proteins are used in the development of a vaccine. Each of these genes is subcloned into a plasmid and transformed into an *E. coli* strain as described above.

The recombinant protein is overexpressed using a vector with a strong promoter. Vectors for use in this technique include pREST (Invitrogen Inc., CA), pKK233-3 (Pharmacia, CA), and the pET system (Promega, WI), although any vector with a strong promoter can be used. After overexpression, the proteins are purified and mixed with adjuvant. Potential adjuvants include, but are not limited to, aluminum hydroxide, QuilA, or Montamide. The purified protein is used as immunogen to vaccinate dogs by a technique being known in the art, and incorporated by reference (Chang *et al.*, 1993c; Chang *et al.*, 1995). Briefly, the individual protein is expressed and purified from *E. coli*. Then, the dogs are injected intramuscularly or subcutaneously with the purified recombinant vaccine and adjuvant. This injection elicits an immune response.

### T Cell Epitope Vaccine

Direct cell cytotoxicity mediated by CD8<sup>+</sup> T lymphocytes (CTL) is the major mechanism of defense against intracellular pathogens. These effector lymphocytes eliminate infected cells by recognizing short peptides associated with MHC class I molecules on the cell surface. Exogenous antigens enter the endosomal pathway and are presented to CD4<sup>+</sup> T cells in association with class II molecules whereas endogenously synthesized antigens are presented to CD8<sup>+</sup> T cells in association with MHC class I molecules. *E. canis* is an intracellular pathogen that resides in monocytes and macrophages. The present invention develops novel ways of generating an *E. canis*-specific CTL response that would eliminate the organism from monocytes or macrophages of infected animals.

A strategy for increasing the protective response of a protein vaccine is to immunize with selective epitopes of the protein. The rationale behind this is that an epitope vaccine contains the most relevant immunogenic peptide components without the irrelevant portions. Therefore, a search is performed for the most highly antigenic portions of the newly identified proteins.

To identify T-cell epitopes from the newly discovered proteins, an initial electronic search for homologous sequences to known T-cell epitopes is performed. In addition, extensive T-cell epitope mapping is carried out. Each of the proteins, ProA, ProB, ORF, the cytochrome oxidase homolog, and the partial lipoprotein signal peptidase homolog, is tested for immunogenic peptide fragments. Mapping of T cell epitopes by a technique known in the art is hereby incorporated by reference (Launois *et al.*, 1994; Lee and Horwitz, 1999). Briefly, short, overlapping peptide sequences (9-20 amino acids) are synthesized over the entire length of the protein in question. These short peptide fragments are tested using healthy dogs which have been immunized with the protein of interest. Peripheral blood mononuclear cells from the dogs are tested for T cell stimulatory and IFN- $\gamma$  inducing properties. Those fragments which elicit the strongest response are the best candidates for a T-cell epitope vaccine.

Once fragments are identified which will make the best epitopes, a recombinant adenylate cyclase of *Bordetella bronchiseptica* is constructed carrying an *E. canis* CD8<sup>+</sup> T cell epitope. The adenylate cyclase toxin (CyaA) of *Bordetella bronchiseptica* causes disease in dogs and elicits an immune response. In addition, CyaA is well suited for intracytoplasmic targeting. Its catalytic domain (AC), corresponding to the N-terminal 400 amino acid residues of the 1,706-residue-long protein, can be delivered to many eukaryotic cells, including cells of the immune system. Also, toxin internalization is independent of receptor-mediated endocytosis, suggesting that the catalytic domain can be delivered directly to the cytosol of target cells through the cytoplasmic membrane. The *Pseudomonas aeruginosa* exotoxin A (PE) is another toxin which could be used in this procedure to deliver peptides or proteins into cells, by a technique known in the art, and hereby incorporated by reference (Donnelly *et al.*, 1993).

Foreign peptides (16 residues) have been inserted into various sites of the AC domain of CyaA without altering its stability or catalytic and calmodulin-binding properties. Thus, protein engineering allows the design and delivery of antigens that specifically stimulate CTLs. The induction of specific CD8<sup>+</sup> T cells can play an important role in canine ehrlichiosis control due to the intracellular persistence of *E. canis* in monocytes.

The adenylate cyclase (AC) toxin (*cya*) gene of *B. bronchiseptica* has been cloned. A synthetic double-stranded oligonucleotide encoding a 9 to 20 amino acid class I T cell epitope of either ProA, ProB, ORF, the cytochrome oxidase homolog, or the partial lipoprotein signal peptidase homolog, is designed according to *B. bronchiseptica* codon usage. The complementary oligonucleotides are inserted in the hypervariable region of the cloned AC-coding sequence of the *cya*. This technique is known in the art in other systems, and is incorporated by reference (Sebo *et al.*, 1995; Guermonprez *et al.*, 1999).

Recombinant plasmids carrying the chimeric *cya* gene are sequenced to determine the copy number and orientation of the inserted epitope. A plasmid with a complete copy of the insert that specifies the T-cell epitope (CD8<sup>+</sup>) in the correct orientation is chosen from the sequenced plasmids. The ability of the new chimeric protein to enter eukaryotic cells is necessary to ensure intracellular targeting of the epitopes (Fayolle *et al.*, 1996).

A vaccine can be created in one of two ways. Recombinant chimeric protein can be purified and used to inoculate dogs. Alternatively, an attenuated *B. bronchiseptica* strain that carries a T-cell epitope or *E. canis* gene by in-frame insertion into adenylate cyclase is created by allelic-exchange. Allelic-exchange is a technique known in the art, and is hereby incorporated by reference (Cotter and Miller, 1994).

Finally, protection against *E. canis* infection in dogs vaccinated with the adenylase cyclase- ProA, ProB, ORF, cytochrome oxidase homolog, or lipoprotein signal peptidase homolog chimeric protein is determined. Wild type and recombinant ACs and CyAs are diluted to working concentrations in PBS and the chimeric protein is injected into dogs either intramuscularly or subcutaneously. Alternatively, the T-cell epitope is inserted into the adenylate cyclase gene of an attenuated *B. bronchiseptica* strain in frame, and the dogs are given the live bacteria.

Recombinant antigens are promising candidates for human and animal vaccination against various pathogens. However, a serious drawback is the poor immunogenicity of recombinant antigens as compared to native antigens. A major challenge in the development of a new recombinant vaccine is, therefore, to have a new adjuvant system that increases the immunogenicity of antigens. Cytokines are powerful immunoregulatory molecules. Cytokines which could be used as adjuvants in this invention include, but are not limited to, IL-12 (interleukin-12), GM-CSF (granulocyte-macrophage colony stimulating factor), IL-1 $\beta$  (interleukin-1 $\beta$ ) and  $\gamma$ -IFN (gamma interferon).

These cytokines can have negative side effects including pyrogenic and/or proinflammatory symptoms in the vaccinated host. Therefore, to avoid the side effects of a whole cytokine protein, an alternate approach is to use synthetic peptide fragments with the desired immunostimulatory properties. The nonapeptide sequence VQGEESNDK of IL-1 $\beta$  protein is endowed with powerful immuno-enhancing properties, and is discussed here to illustrate the use of a cytokine to increase immunogenicity.

This nonapeptide is inserted into the ProA, ProB, ORF, the cytochrome oxidase homolog, or the partial lipoprotein signal peptidase homolog protein and its immunogenicity is compared to that of the native protein. Reportedly, the insertion of this

sequence into a poorly immunogenic recombinant antigen increases the chance of a strong protective immune response after vaccination. This peptide could enhance the *in vivo* immune response against both T-dependent and T-independent antigens. The canine IL-1 $\beta$  sequence may mimic many immunomodulatory activities of the entire molecule of IL-1 $\beta$  while apparently lacking many of its undesirable proinflammatory properties. This strategy is employed to increase the immunogenicity of ProA, ProB, ORF, cytochrome oxidase, the partial lipoprotein signal peptidase homolog and other *E. canis* antigens.

Plasmid pYFC199 is derived from a pBR322 plasmid by the insertion of a fragment that includes the ProA, ProB, ORF, the cytochrome oxidase homolog, or the partial lipoprotein signal peptidase protein from *E. canis*. This plasmid contains a unique *Hind*III site where in-frame insertions encoding exogenous sequences can be inserted. Two complementary oligonucleotides, AGGCTTGTTTCAGGGTGAAGAAGAATCCAACGACAAAAGCTT and AAGCTTTTGTCTGTTGGATTCTTCACCCTGAACTTGCCA, that encode the canine IL-1 $\beta$  163-171 peptide are annealed, cut with *Hind*III, and inserted into the pYFC199 *Hind*III site. The recombinant plasmid carrying the chimeric IL-1 $\beta$  gene is sequenced to determine the orientation of the inserted epitope.

The efficacy of the recombinant proteins as vaccines is tested in dogs. The purified protein is injected intraperitoneally into dogs. Specific pathogen free (SPF) dogs are divided into five groups: one group is given recombinant adenylate cyclase of *Bordetella bronchiseptica* carrying *E. canis* CD8<sup>+</sup> T cell epitopes derived from ProA, ProB, ORF, cytochrome oxidase homolog, or the partial lipoprotein signal peptidase homolog, one group is given recombinant adenylate cyclase of *Bordetella bronchiseptica* as a control, one group is given the ProA, ProB, ORF, cytochrome oxidase homolog, or the partial lipoprotein signal peptidase homolog protein plus a canine IL-1 $\beta$  163-171 insert, one group is given a T cell epitope derived from ProA, ProB, ORF, cytochrome oxidase homolog, or the partial lipoprotein signal peptidase homolog alone, and the last group is given PBS as a negative control.

All animals are vaccinated (30-40  $\mu$ g each) four times. The dogs are challenged ten days after the last vaccination with  $10^7$  *E. canis*. At day five postchallenge, approximately 1 ml blood from each dog is collected in an EDTA tube. Whether the vaccinated groups eliminate the organisms as compared to that of the control group is tested by culture and PCR. Two primers derived from the genes cloned can be used to amplify the gene product from the tissues or blood samples from these dogs. The internal primer can also be designed for use as an oligonucleotide probe to hybridize the PCR gene product.

This invention provides a badly needed vaccine against the *E. canis* bacterium. The vaccine can be used to protect dogs throughout the world from canine monocytic ehrlichiosis.

Accordingly, it is to be understood that the embodiments of the invention herein described are merely illustrative of the application of the principles of the invention. Reference herein to details of the illustrated embodiments are not intended to limit the scope of the claims, which themselves recite those features regarded as essential to the invention.

What is claimed is:

- 1 1. A recombinant DNA comprising said DNA selected from the group consisting of:
  - 2 a) a recombinant DNA that encodes a protein having an amino acid sequence as  
3 shown in SEQ. ID. NO. 3;
  - 4 b) a recombinant DNA that encodes a protein having an amino acid sequence as  
5 shown in SEQ. ID. NO. 5;
  - 6 c) a recombinant DNA that encodes a protein having an amino acid sequence as  
7 shown in SEQ. ID. NO. 7;
  - 8 d) a recombinant DNA that encodes a protein having an amino acid sequence as  
9 shown in SEQ. ID. NO. 9;
  - 10 e) a recombinant DNA that encodes a protein having an amino acid sequence as  
11 shown in SEQ. ID. NO. 11; and
  - 12 f) any portion of said DNA above that encodes a protein that elicits an immune  
13 response against *E. canis*.
- 1 2. The recombinant DNA of claim 1 wherein said DNA encodes at least one  
2 immunogenic epitope.
- 1 3. A recombinant protein comprising said protein selected from the group consisting of:
  - 2 a) a protein having an amino acid sequence as shown in SEQ. ID. NO. 3;
  - 3 b) a protein having an amino acid sequence as shown in SEQ. ID. NO. 5;
  - 4 c) a protein having an amino acid sequence as shown in SEQ. ID. NO. 7;
  - 5 d) a protein having an amino acid sequence as shown in SEQ. ID. NO. 9;
  - 6 e) a protein having an amino acid sequence as shown in SEQ. ID. NO. 11; and

7 f) any portion of any of the above proteins that elicits an immune response  
8 against *E. canis*.

1 4. The recombinant protein of claim 3 wherein said protein includes at least one  
2 immunogenic epitope.

1 5. A vaccine wherein said vaccine protects dogs against *E. canis* infection.

1 6. The vaccine of claim 5 comprising:

2 a) a vector capable of expressing a recombinant DNA inserted into said vector  
3 such that a recombinant protein is expressed when said vector is provided in an  
4 appropriate host; and

5 b) the recombinant DNA inserted into said vector wherein said DNA is selected  
6 from the group consisting of:

7 i. a recombinant DNA that encodes a protein having an amino acid  
8 sequence as shown in SEQ. ID. NO. 3;

9 ii. a recombinant DNA that encodes a protein having an amino acid  
10 sequence as shown in SEQ. ID. NO. 5;

11 iii. a recombinant DNA that encodes a protein having an amino acid  
12 sequence as shown in SEQ. ID. NO. 7;

13 iv. a recombinant DNA that encodes a protein having an amino acid  
14 sequence as shown in SEQ. ID. NO. 9;

15 v. a recombinant DNA that encodes a protein having an amino acid  
16 sequence as shown in SEQ. ID. NO. 11; and

17 vi. any portion of said DNA above that encodes a protein that elicits an  
18 immune response against *E. canis*.



1 7. The vaccine of claim 6, wherein said DNA further comprises DNA that encodes CpG  
2 motifs.

1 8. The vaccine of claim 6 wherein said DNA further comprises a promoter selected from  
2 the group consisting of:

3 a) a cytomegalovirus (CMV) immediate early promoter;

4 b) a human tissue plasminogen activator gene (t-PA); and

5 c) a promoter/enhancer region of a human elongation factor alpha (EF-1  $\alpha$ ).

1 9. The vaccine of claim 6, wherein said vector is selected from the group consisting of:

2 a) pcDNA3;

3 b) pC1;

4 c) VR1012; and

5 d) VR1020.

1 10. The vaccine of claim 6 wherein said vaccine is administered into said host by a method  
2 selected from the group consisting of:

3 a) intramuscular injection;

4 b) intravenous injection; and

5 c) gene gun injection.

1 11. The vaccine of claim 10, wherein said host is a dog.

1 12. The vaccine of claim 5 comprising:

2 a) a recombinant protein that is selected from the group consisting of:

3 i. a protein having an amino acid sequence as shown in SEQ. ID. NO. 3;

- 4           ii. a protein having an amino acid sequence as shown in SEQ. ID. NO. 5;
- 5           iii. a protein having an amino acid sequence as shown in SEQ. ID. NO. 7;
- 6           iv. a protein having an amino acid sequence as shown in SEQ. ID. NO. 9;
- 7           v. a protein having an amino acid sequence as shown in SEQ. ID. NO. 11;
- 8           and
- 9           vi. any portion of any of the above proteins that elicits an immune response
- 10           against *E. canis*.

1       13. The vaccine of claim 12, wherein said vaccine further comprises adjuvants selected

2       from the group consisting of:

- 3           a) aluminum hydroxide;
- 4           b) QuilA; and
- 5           c) Montamide.

1       14. The vaccine of claim 12 further comprising a cytokine operatively associated with said

2       recombinant protein.

1       15. The vaccine of claim 14 wherein said cytokine is selected from the group consisting

2       of:

- 3           a) interleukin-1 $\beta$  (IL-1 $\beta$ );
- 4           b) granulocyte-macrophage colony stimulating factor (GM-CSF);
- 5           c) gamma interferon ( $\gamma$ -IFN);
- 6           d) amino acids VQGEESNDK from the IL-1 $\beta$  protein; and
- 7           e) any portion of any of the cytokines above that elicits an improved
- 8           immunogenic response against *E. canis*.

1 16. The vaccine of claim 12 wherein said vaccine is administered into a host by a method  
2 selected from the group consisting of:

- 3 a) intramuscular injection; and  
4 b) subcutaneous injection.

1 17. The vaccine of claim 16 wherein said host is a dog.

1 18. The vaccine of claim 5 comprising a recombinant protein that includes a T cell epitope  
2 wherein said T cell epitope comprises an amino acid peptide fragment of a protein  
3 selected from the group consisting of:

- 4 a) a protein having an amino acid sequence as shown in SEQ. ID. NO. 3;  
5 b) a protein having an amino acid sequence as shown in SEQ. ID. NO. 5;  
6 c) a protein having an amino acid sequence as shown in SEQ. ID. NO. 7;  
7 d) a protein having an amino acid sequence as shown in SEQ. ID. NO. 9;  
8 e) a protein having an amino acid sequence as shown in SEQ. ID. NO. 11; and  
9 f) any portion of any of the above proteins that elicits an immune response  
10 against *E. canis*.

1 19. The vaccine of claim 18 wherein said amino acid peptide fragment comprises nine to  
2 twenty amino acids.

1 20. The vaccine of claim 18 further comprising a recombinant DNA encoding a protein  
2 which is capable of being internalized into eukaryotic cells, including cells of the  
3 immune system.

1 21. The vaccine of claim 20 wherein said protein capable of being internalized into  
2 eukaryotic cells comprises a toxin selected from the group consisting of:

- 3 a) a recombinant adenylate cyclase of *Bordetella bronchiseptica*; and

4           b) a recombinant exotoxin A (PE) of *Pseudomonas aeruginosa*.

1       22. The vaccine of claim 18 wherein said vaccine is administered into a host by a method  
2       selected from the group consisting of:

3           a) intramuscular injection; and

4           b) subcutaneous injection.

1       23. The vaccine of claim 22 wherein said host is a dog.

1       24. A method of identifying a T cell epitope against *E. canis* comprising:

2           a) synthesizing overlapping peptide fragments over an entire length of a protein  
3           wherein said protein is selected from the group consisting of:

4               i. a protein having an amino acid sequence as shown in SEQ. ID. NO. 3;

5               ii. a protein having an amino acid sequence as shown in SEQ. ID. NO. 5;

6               iii. a protein having an amino acid sequence as shown in SEQ. ID. NO. 7;

7               iv. a protein having an amino acid sequence as shown in SEQ. ID. NO. 9;

8               v. a protein having an amino acid sequence as shown in SEQ. ID. NO. 11;  
9               and

10              vi. any portion of any of the proteins above that elicits an immune response  
11              against *E. canis*;

12           b) testing said peptide fragment to determine if said peptide fragment elicits an  
13           immune response in a host animal; and

14           c) identifying said peptide fragment as said T cell epitope of *E. canis* if said  
15           fragment elicits an immune response.

1       25. The method of claim 24 wherein said peptide fragment comprises nine to twenty  
2       amino acids.

1 26. A method of creating a vaccine against *E. canis* comprising:

2 a) selecting a vector capable of expressing a recombinant DNA inserted into said  
3 vector; and

4 b) inserting a recombinant DNA into said vector such that a recombinant protein  
5 is expressed when said vector is provided in an appropriate host wherein said  
6 DNA is selected from the group consisting of:

7 i. a recombinant DNA that encodes a protein having an amino acid  
8 sequence as shown in SEQ. ID. NO. 3;

9 ii. a recombinant DNA that encodes a protein having an amino acid  
10 sequence as shown in SEQ. ID. NO. 5;

11 iii. a recombinant DNA that encodes a protein having an amino acid  
12 sequence as shown in SEQ. ID. NO. 7;

13 iv. a recombinant DNA that encodes a protein having an amino acid  
14 sequence as shown in SEQ. ID. NO. 9;

15 v. a recombinant DNA that encodes a protein having an amino acid  
16 sequence as shown in SEQ. ID. NO. 11; and

17 vi. any portion of said DNA above that encodes a protein that elicits an  
18 immune response against *E. canis*.

1 27. The method of claim 26, wherein said DNA further comprises DNA that encodes CpG  
2 motifs.

1 28. The method of claim 26 wherein said DNA further comprises a promoter selected from  
2 the group consisting of:

3 a) a cytomegalovirus (CMV) immediate early promoter;

4 b) a human tissue plasminogen activator gene (t-PA); and

5 c) a promoter/enhancer region of a human elongation factor alpha (EF-1  $\alpha$ ).

1 29. The method of claim 26, wherein said vector is selected from the group consisting of:

2 a) pcDNA3;

3 b) pC1;

4 c) VR1012; and

5 d) VR1020.

1 30. The method of claim 26 wherein said vaccine is injected into said host in a manner  
2 selected from the group consisting of:

3 a) intramuscular injection;

4 b) intravenous injection; and

5 c) gene gun injection.

1 31. The method of claim 30, wherein said host is a dog.

1 32. A method of creating a vaccine against *E. canis* comprising:

2 a) selecting a vector capable of expressing a recombinant protein inserted into  
3 said vector;

4 b) insertion of a recombinant DNA into said vector such that said recombinant  
5 protein is expressed when said vector is transformed into a bacterial strain  
6 wherein said DNA is selected from the group consisting of:

7 i. a recombinant DNA that encodes a protein having an amino acid  
8 sequence as shown in SEQ. ID. NO. 3;

9 ii. a recombinant DNA that encodes a protein having an amino acid  
10 sequence as shown in SEQ. ID. NO. 5;

- 11           iii. a recombinant DNA that encodes a protein having an amino acid  
12           sequence as shown in SEQ. ID. NO. 7;
- 13           iv. a recombinant DNA that encodes a protein having an amino acid  
14           sequence as shown in SEQ. ID. NO. 9;
- 15           v. a recombinant DNA that encodes a protein having an amino acid  
16           sequence as shown in SEQ. ID. NO. 11; and
- 17           vi. any portion of said DNA above that encodes a protein that elicits an  
18           immune response against *E. canis*; and
- 19           c) harvesting said recombinant protein from said bacterial strain.

1   33. The method of claim 32, wherein said vaccine further comprises adjuvants selected  
2   from the group consisting of:

- 3           a) aluminum hydroxide;
- 4           b) QuilA; and
- 5           c) Montamide.

1   34. The method of claim 32, wherein said vaccine further comprises a promoter selected  
2   from the group consisting of:

- 3           a) tac;
- 4           b) T5; and
- 5           c) T7.

1   35. The method of claim 32, wherein said bacterial strain is *E. coli*.

1   36. The method of claim 32, wherein said vector is selected from the group consisting of:

- 2           a) pREST;

3           b) pET; and

4           c) pKK233-3.

1       37. The method of claim 32 wherein said vaccine further comprises a cytokine operatively  
2       associated with said vaccine.

1       38. The method of claim 37 wherein said cytokine is selected from the group consisting  
2       of:

3           a) interleukin-1 $\beta$  (IL-1 $\beta$ );

4           b) granulocyte-macrophage colony stimulating factor (GM-CSF);

5           c) gamma interferon ( $\gamma$ -IFN);

6           d) amino acids VQGEESNDK from the IL-I $\beta$  protein; and

7           e) any portion of any of the cytokines above that elicits an improved  
8       immunogenic response against *E. canis*.

1       39. The method of claim 32 wherein said vaccine is injected into said host in a manner  
2       selected from the group consisting of:

3           a) intramuscular injection; and

4           b) subcutaneous injection.

1       40. The method of claim 39 wherein said host is a dog.

1       41. A method of creating a T cell epitope vaccine comprising:

2           a) selecting a recombinant protein that includes a T cell epitope wherein said T  
3       cell epitope comprises an amino acid peptide fragment of a protein selected  
4       from the group consisting of:

5           i. a protein having an amino acid sequence as shown in SEQ. ID. NO. 3;



- ii. a protein having an amino acid sequence as shown in SEQ. ID. NO. 5;
- iii. a protein having an amino acid sequence as shown in SEQ. ID. NO. 7;
- iv. a protein having an amino acid sequence as shown in SEQ. ID. NO. 9;
- v. a protein having an amino acid sequence as shown in SEQ. ID. NO. 11;
- and
- vi. any portion of any of the above proteins that elicits an immune response against *E. canis*;

b) identifying said T cell epitope from said protein;

c) incorporating said T cell epitope into a construct capable of expressing said epitope as a protein; and

d) harvesting said protein.

42. The method of claim 41 wherein said amino acid peptide fragment comprises nine to twenty amino acids.

43. The method of claim 41 wherein said construct capable of expressing said epitope further comprises a recombinant DNA encoding a protein which is capable of being internalized into eukaryotic cells, including cells of the immune system.

44. The method of claim 43 wherein said protein capable of being internalized into eukaryotic cells comprises a toxin selected from the group consisting of:

a) a recombinant adenylate cyclase of *Bordetella bronchiseptica*; and

b) a recombinant exotoxin A (PE) of *Pseudomonas aeruginosa*.

45. The method of claim 41 wherein said vaccine is injected into said host in a manner selected from the group consisting of:

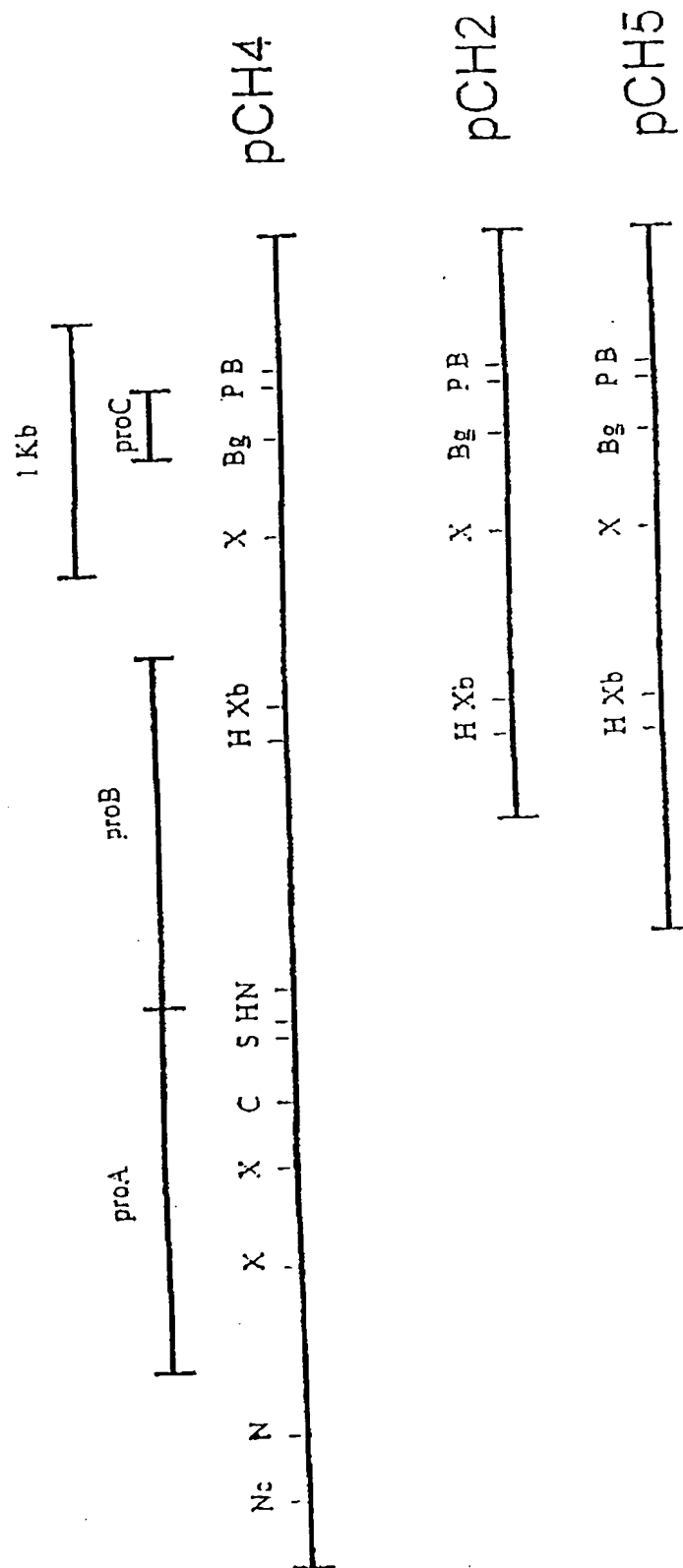
a) intramuscular injection; and

4           b) subcutaneous injection.

1       46. The method of claim 45 wherein said host is a dog.

1/1

Figure 1



## SEQUENCE LISTING

<110> Chang, Yung-Fu

<120> EHRLICHIA CANIS GENES FOR VACCINE DEVELOPMENT

<130> crf2322

<150> U.S. 09/358,322

<151> 1999-07-21

<160> 11

<170> PatentIn Ver. 2.0

<210> 1

<211> 5300

<212> DNA

<213> Ehrlichia canis

<400> 1

gatcaaataa aatgaaacca agaataagaa acactattta tggattaata gcaataatac 60  
tatctatgat atgttttagtg tacgcttctg taccactata tagtatattt tgtaaagtaa 120  
caggttatgg aggtacagta agaacaagta atatatcaaa ttctaaaata ggtaacacta 180  
ttattaaagt cagatttaat gcagatatac acaacaact gccatggaaa ttctatccag 240  
aagtatctca tgtatttgta aaaccaggag aacaaaaatt gattttctac cgcgcagaaa 300  
atctacttga tgaggacact tcaggaatgg ctgtatataa tgttacacca cataaagtag 360  
gaaaatattt taataaggta gcttggtttt gtttcaccaa acaaacatta taccctcatc 420  
aaaaaactat aatgccagta tcatttttta tagatccagc catagaaaca gatcctgaaa 480  
ctgctgacgt aaaactcatc actctttcat atgtattctt taagtacaaa gaataaactt 540  
catataccgt acattataaa ctgattaaaa aaaataacta ttaatatga gcaaaataat 600  
ttatctattc aacagattct tttcaattag agagtattca aaaacactac aactactgct 660  
tgcaactttc tatcactgat atataaaagt gaaataaatt taaaaaactt tagttttaat 720  
agaagaattt tattaaaagc tttgaatcaa atttaattac tgatataaaa atactattaa 780  
acattaacaa tgcttaatta aagtattatt atttacctta atttcataac ctttattaac 840

aatttcataa taaaaatact ttactcttat ttttttatca cttgatatta ttaaataatc 900  
atataaactc ccaaataaac tattgcaagg ttatggtaat gatgaaattt ttacttggtt 960  
ttttcatagt tttcttaaca atagccaatc atgctttatc ctttaacatt aaagttacac 1020  
atgaaaaatt agataatgga atggaagtat acgtgattcc aaatcatcgc gcaccagcag 1080  
tcatgcacat ggtattatac aaagtcggtg gaactgatga tccagtagga tactctggat 1140  
tagcacattt ttttgaacac ttaatgttta gtggaacaga aaaatttcct aatctcatca 1200  
gcacacttag taatataggc ggaaatttca atgcaagcac atctcaattt tgtactatat 1260  
actacgaatt aataccaaaa caatatattat ctcttgcaat ggatattgaa tcagacagaa 1320  
tgcagaattt taaggttacc gacaaagcat taataagaga acaaaaggta gtcttagaag 1380  
aaagaaaaat gagagttgaa agccaagcaa aaaacatact agaagaagaa atggaaaatg 1440  
cattttatta caatggatat ggcagaccag tagtaggatg ggaacatgaa attagcaact 1500  
acaacaaaga agttgctgaa gcctttcata agctacatta tagtcctaata aatgctatat 1560  
taattgtaac tggagatgca gatccacaag aagtaatcac acttgcaaaa caatactatg 1620  
ggaaaatacc atctaataat aagaaacctt caagtcaagt tagggtagaa ccaccgcata 1680  
aaacaaatat gactttaaca ttaaaagaca gttcagtaga aatcccagaa ctgtttttta 1740  
tgtatcaaat accaaatggg attaccaata aaaactacat acttaacatg atggttagcag 1800  
aaatactcgg tagtggtaaa ttcagcctgc tttacaatga tttggtaatt aacaatccaa 1860  
tagttacatc gataaaaaca gattataatt acttaactga cagcgataat tacctttcca 1920  
ttgaagctat acctaaaaac gggatctcta cagaagctgt agaacaagaa attcataaat 1980  
gtataaataa ttatttagaa aatggaattt cagcagaata tttagaaagt gcaaagtata 2040  
aagtaaaagc acatttaact tatgcatttg acggactaac tttcatatca tatttttatg 2100  
gcatgcctct aatactagga gtaccgctat cagaaatcag taatatttac gataccatag 2160  
acaaagtaag tatccaagat gtttaactccg ctatggaaaa tatctttcaa aacaatataa 2220  
gattaaccgg gcatttatta cctaattggag aatagttatg agaaacatat tgtgttacac 2280  
attaatattg attttctttt cattcaatac atatgcaaat gatctcaata ttaacataaa 2340  
agaagctaca actaaaaata aaatacacta tctatatgtt gaacatcata acctaccaac 2400  
aatttcctta aaatttgcat tcaagaaagc aggatacgct tatgatgcct ttgataagca 2460  
aggacttgca tactttacat caaaaatatt aaacgaagga tcaaaaaaca actatgctct 2520  
cagttttgca caacaattag aaggcaaagg tatagactta aaatttgata tagacctaga 2580

caatTTTTat atatcattaa aaaccttatac agaaaacttt gaagaagccc tagttttact 2640  
cagtgattgc atattcaaca ccgtcacaga tcaagaaata ttcaatagaa taatagcaga 2700  
acagattgca catgttaa at cattatatc tgctcctgaa tttatagcta caacagaaat 2760  
gaatcacgct atattcaaag ggcacccata ttctaacaaa gtttacggga cattaaatac 2820  
aatcaataat atcaaccagg aagacgttgc attatatata aaaaatagtt ttgacaagga 2880  
acaaatcggt atcagcgcag caggagatgt agatccaaca cagctatcaa atttactaga 2940  
taaataatatt ctttccaaat tgccatctgg taataacaaa aataccatac cagatacgac 3000  
tgттаataga gaagacacat tattatatgt acagagagat gtaccacaaa gtgtcataat 3060  
gtttgctaca gacacagtac catatcacag caaagactat catgcatcaa acttgttcaa 3120  
tactatgcta ggcggattaa gtctcaattc aatattaatg atagaattaa gagacaagtt 3180  
aggattaaca taccatagta gcagttcact atctaacatg aatcatagta atgtgctatt 3240  
tggtacaata ttactgata ataccacagt aacaaaatgt atatccgtct taacagatat 3300  
tatagagcac attaaaaagt atggagttga tgaagacact tttgcaattg caaaatctag 3360  
tattaccaac tcttttattt tatctatggt aaataacaat aatgttagtg agatattggt 3420  
aagcttaca ttacacgatac tagatccgag ttatattaat aaatacaatt cttactacaa 3480  
agcaataaca atagaagaag taaataaaat tgccaagaaa attttatcta atgaattagt 3540  
aataattgaa gtaggaaaaa acaataacat aaatggcaaa caaatagatg ctaaaaaaca 3600  
cataccttgg ttaagtatac aggttattgt atttactaca agtattctat taggttgtat 3660  
taagtaagta taagtagctt caatcaaata aaaaaacatt aaccaaagtg ttagctctac 3720  
cggagaagct tattataagc ttttaacctg ggataatatg aagttttgct aatgttaagc 3780  
aaaaaattag taatcacaat atcaaatttt ctttacagga ttatattgtg acctaccata 3840  
acaacttata tttagaaaat gacaacagat acacacatca ataaattatc actacaattc 3900  
aattaataaa acaatgagta tttttactta attatttaat tttatttttt aaaataaaat 3960  
tacaatttta cttactcaat aaaagcagtt atactaccaa gtattggatg gtattaatcg 4020  
gagcaattac tacttaatag tatagctggt gacaagccgc aatctgcggt tcttgacaaa 4080  
ataatactaa tcagttaaaa ttttgaagtg tttcaccata atgggtattat ttatgaaagc 4140  
tcatagcaca agtatagga actttcagcc tttagaaaga gctgctataa tcattgcagt 4200  
gttaggttta gctgcattct tgtttgctgc tgctgcctgc agtgatcggt tccaaagatt 4260

gcaattaaca aatccatttg taatagcagg aatggttggc cttgcagttc ttttagttgc 4320  
ttccttaaca gcagcattaa gtatatgctt aactaaaagt aagcaagtca cacaacatgc 4380  
tattagacat cgctttggat acgagtcaag cacttcttct tctgtactgc ttgcaatata 4440  
aataatttct ttattacttg ctgcagcatt ttgtggaaag ataatgggta atgacaaccc 4500  
agatctattc tttagcaaga tgcaagaact ctccaatcca cttgttgttg cagctattgt 4560  
agccgtttct gttttcctac tctcatctgt aatgtatgct gcaaagaaca ttataagtcc 4620  
agataaaca actcacgtta ttatattatc taatcaaca actatagaag aagcaaaagt 4680  
agatcaagga atgaatattt tgtcagcagt actcccagca gctggcattg acatcatgac 4740  
tatagcttct tgtgacattt tagcagtgag cagccgggga tcctctcagc atcaatagat 4800  
ttatgtttta gcctgtattc acctttttat taggtgttgt atcgtttctt tatataagtg 4860  
tgttatatta tataaaacat ctaggagtta cagttaattt gtttcatgtg gttattactc 4920  
tttgccatta ttattactat acctaaaaat ataaaagaat ccgccagggt gaatacaggc 4980  
caatgtaagt tattgatata aaaatctata aaatcataga cagcaccata tcttattcta 5040  
tctatgatat ttcctattga ccccccaata atgattacaa gaggtaatct ataatgtggc 5100  
tgtactataa ataagtagca taaaacacaa gtaatcaaaa tcgagatact acaaaaaaca 5160  
acattactat attcaaagtt atttaataata caaaactaa ttccagcatt ccacactgta 5220  
gtaaagcgca agaagcttaa tatctctatt acacctttat ctctatcaa atttactaca 5280  
taccatttac ttacctgac 5300

<210> 2

<211> 522

<212> DNA

<213> Ehrlichia canis

<220>

<221> CDS

<222> (1)..(522)

<223> Protein translated from nucleotides 12 through 533  
(cytochrome oxidase homolog).

<400> 2

5

atg aaa cca aga ata aga aac act att tat gga tta ata gca ata ata 48  
 Met Lys Pro Arg Ile Arg Asn Thr Ile Tyr Gly Leu Ile Ala Ile Ile  
 1 5 10 15  
 cta tct atg ata tgt tta gtg tac gct tct gta cca cta tat agt ata 96  
 Leu Ser Met Ile Cys Leu Val Tyr Ala Ser Val Pro Leu Tyr Ser Ile  
 20 25 30  
 ttt tgt aaa gta aca ggt tat gga ggt aca gta aga aca agt aat ata 144  
 Phe Cys Lys Val Thr Gly Tyr Gly Gly Thr Val Arg Thr Ser Asn Ile  
 35 40 45  
 tca aat tct aaa ata ggt aac act att att aaa gtc aga ttt aat gca 192  
 Ser Asn Ser Lys Ile Gly Asn Thr Ile Ile Lys Val Arg Phe Asn Ala  
 50 55 60  
 gat ata cac aaa caa ctg cca tgg aaa ttc tat cca gaa gta tct cat 240  
 Asp Ile His Lys Gln Leu Pro Trp Lys Phe Tyr Pro Glu Val Ser His  
 65 70 75 80  
 gta ttt gta aaa cca gga gaa caa aaa ttg att ttc tac cgc gca gaa 288  
 Val Phe Val Lys Pro Gly Glu Gln Lys Leu Ile Phe Tyr Arg Ala Glu  
 85 90 95  
 aat cta ctt gat gag gac act tca gga atg gct gta tat aat gtt aca 336  
 Asn Leu Leu Asp Glu Asp Thr Ser Gly Met Ala Val Tyr Asn Val Thr  
 100 105 110  
 cca cat aaa gta gga aaa tat ttt aat aag gta gct tgt ttt tgt ttc 384  
 Pro His Lys Val Gly Lys Tyr Phe Asn Lys Val Ala Cys Phe Cys Phe  
 115 120 125  
 acc aaa caa aca tta tac cct cat caa aaa act ata atg cca gta tca 432  
 Thr Lys Gln Thr Leu Tyr Pro His Gln Lys Thr Ile Met Pro Val Ser  
 130 135 140  
 ttt ttt ata gat cca gcc ata gaa aca gat cct gaa act gct gac gta 480  
 Phe Phe Ile Asp Pro Ala Ile Glu Thr Asp Pro Glu Thr Ala Asp Val  
 145 150 155 160  
 aaa ctc atc act ctt tca tat gta ttc ttt aag tac aaa gaa 522  
 Lys Leu Ile Thr Leu Ser Tyr Val Phe Phe Lys Tyr Lys Glu  
 165 170

&lt;210&gt; 3

&lt;211&gt; 174

&lt;212&gt; PRT

&lt;213&gt; Ehrlichia canis

&lt;400&gt; 3

Met Lys Pro Arg Ile Arg Asn Thr Ile Tyr Gly Leu Ile Ala Ile Ile  
 1 5 10 15



6

Leu Ser Met Ile Cys Leu Val Tyr Ala Ser Val Pro Leu Tyr Ser Ile  
                   20                  25                  30  
 Phe Cys Lys Val Thr Gly Tyr Gly Gly Thr Val Arg Thr Ser Asn Ile  
                   35                  40                  45  
 Ser Asn Ser Lys Ile Gly Asn Thr Ile Ile Lys Val Arg Phe Asn Ala  
                   50                  55                  60  
 Asp Ile His Lys Gln Leu Pro Trp Lys Phe Tyr Pro Glu Val Ser His  
                   65                  70                  75                  80  
 Val Phe Val Lys Pro Gly Glu Gln Lys Leu Ile Phe Tyr Arg Ala Glu  
                   85                  90                  95  
 Asn Leu Leu Asp Glu Asp Thr Ser Gly Met Ala Val Tyr Asn Val Thr  
                   100                  105                  110  
 Pro His Lys Val Gly Lys Tyr Phe Asn Lys Val Ala Cys Phe Cys Phe  
                   115                  120                  125  
 Thr Lys Gln Thr Leu Tyr Pro His Gln Lys Thr Ile Met Pro Val Ser  
                   130                  135                  140  
 Phe Phe Ile Asp Pro Ala Ile Glu Thr Asp Pro Glu Thr Ala Asp Val  
                   145                  150                  155                  160  
 Lys Leu Ile Thr Leu Ser Tyr Val Phe Phe Lys Tyr Lys Glu  
                   165                  170

&lt;210&gt; 4

&lt;211&gt; 1314

&lt;212&gt; DNA

&lt;213&gt; Ehrlichia canis

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1314)

<223> Protein translated from nucleotides 939 through  
 2,252 (ProA).

&lt;400&gt; 4

atg atg aaa ttt ttt act tgt ttt ttc ata gtt ttc tta aca ata gcc 48  
 Met Met Lys Phe Phe Thr Cys Phe Phe Ile Val Phe Leu Thr Ile Ala  
           1                  5                  10                  15

aat cat gct tta tcc ttt aac att aaa gtt aca cat gaa aaa tta gat	96
Asn His Ala Leu Ser Phe Asn Ile Lys Val Thr His Glu Lys Leu Asp	
20 25 30	
aat gga atg gaa gta tac gtg att cca aat cat cgc gca cca gca gtc	144
Asn Gly Met Glu Val Tyr Val Ile Pro Asn His Arg Ala Pro Ala Val	
35 40 45	
atg cac atg gta tta tac aaa gtc ggt gga act gat gat cca gta gga	192
Met His Met Val Leu Tyr Lys Val Gly Gly Thr Asp Asp Pro Val Gly	
50 55 60	
tac tct gga tta gca cat ttt ttt gaa cac tta atg ttt agt gga aca	240
Tyr Ser Gly Leu Ala His Phe Phe Glu His Leu Met Phe Ser Gly Thr	
65 70 75 80	
gaa aaa ttt cct aat ctc atc agc aca ctt agt aat ata ggc gga aat	288
Glu Lys Phe Pro Asn Leu Ile Ser Thr Leu Ser Asn Ile Gly Gly Asn	
85 90 95	
ttc aat gca agc aca tct caa ttt tgt act ata tac tac gaa tta ata	336
Phe Asn Ala Ser Thr Ser Gln Phe Cys Thr Ile Tyr Tyr Glu Leu Ile	
100 105 110	
cca aaa caa tat tta tct ctt gca atg gat att gaa tca gac aga atg	384
Pro Lys Gln Tyr Leu Ser Leu Ala Met Asp Ile Glu Ser Asp Arg Met	
115 120 125	
cag aat ttt aag gtt acc gac aaa gca tta ata aga gaa caa aag gta	432
Gln Asn Phe Lys Val Thr Asp Lys Ala Leu Ile Arg Glu Gln Lys Val	
130 135 140	
gtc tta gaa gaa aga aaa atg aga gtt gaa agc caa gca aaa aac ata	480
Val Leu Glu Glu Arg Lys Met Arg Val Glu Ser Gln Ala Lys Asn Ile	
145 150 155 160	
cta gaa gaa gaa atg gaa aat gca ttt tat tac aat gga tat ggc aga	528
Leu Glu Glu Glu Met Glu Asn Ala Phe Tyr Tyr Asn Gly Tyr Gly Arg	
165 170 175	
cca gta gta gga tgg gaa cat gaa att agc aac tac aac aaa gaa gtt	576
Pro Val Val Gly Trp Glu His Glu Ile Ser Asn Tyr Asn Lys Glu Val	
180 185 190	
gct gaa gcc ttt cat aag cta cat tat agt cct aat aat gct ata tta	624
Ala Glu Ala Phe His Lys Leu His Tyr Ser Pro Asn Asn Ala Ile Leu	
195 200 205	
att gta act gga gat gca gat cca caa gaa gta atc aca ctt gca aaa	672
Ile Val Thr Gly Asp Ala Asp Pro Gln Glu Val Ile Thr Leu Ala Lys	
210 215 220	
caa tac tat ggg aaa ata cca tct aat aat aag aaa cct tca agt caa	720
Gln Tyr Tyr Gly Lys Ile Pro Ser Asn Asn Lys Lys Pro Ser Ser Gln	
225 230 235 240	

gtt agg gta gaa cca ccg cat aaa aca aat atg act tta aca tta aaa 768  
 Val Arg Val Glu Pro Pro His Lys Thr Asn Met Thr Leu Thr Leu Lys  
 245 250 255

gac agt tca gta gaa atc cca gaa ctg ttt tta atg tat caa ata cca 816  
 Asp Ser Ser Val Glu Ile Pro Glu Leu Phe Leu Met Tyr Gln Ile Pro  
 260 265 270

aat ggt att acc aat aaa aac tac ata ctt aac atg atg tta gca gaa 864  
 Asn Gly Ile Thr Asn Lys Asn Tyr Ile Leu Asn Met Met Leu Ala Glu  
 275 280 285

ata ctc ggt agt ggt aaa ttc agc ctg ctt tac aat gat ttg gta att 912  
 Ile Leu Gly Ser Gly Lys Phe Ser Leu Leu Tyr Asn Asp Leu Val Ile  
 290 295 300

aac aat cca ata gtt aca tcg ata aaa aca gat tat aat tac tta act 960  
 Asn Asn Pro Ile Val Thr Ser Ile Lys Thr Asp Tyr Asn Tyr Leu Thr  
 305 310 315 320

gac agc gat aat tac ctt tcc att gaa gct ata cct aaa aac ggg atc 1008  
 Asp Ser Asp Asn Tyr Leu Ser Ile Glu Ala Ile Pro Lys Asn Gly Ile  
 325 330 335

tct aca gaa gct gta gaa caa gaa att cat aaa tgt ata aat aat tat 1056  
 Ser Thr Glu Ala Val Glu Gln Glu Ile His Lys Cys Ile Asn Asn Tyr  
 340 345 350

tta gaa aat gga att tca gca gaa tat tta gaa agt gca aag tat aaa 1104  
 Leu Glu Asn Gly Ile Ser Ala Glu Tyr Leu Glu Ser Ala Lys Tyr Lys  
 355 360 365

gta aaa gca cat tta act tat gca ttt gac gga cta act ttc ata tca 1152  
 Val Lys Ala His Leu Thr Tyr Ala Phe Asp Gly Leu Thr Phe Ile Ser  
 370 375 380

tat ttt tat ggc atg cat cta ata cta gga gta ccg cta tca gaa atc 1200  
 Tyr Phe Tyr Gly Met His Leu Ile Leu Gly Val Pro Leu Ser Glu Ile  
 385 390 395 400

agt aat att tac gat acc ata gac aaa gta agt atc caa gat gtt aac 1248  
 Ser Asn Ile Tyr Asp Thr Ile Asp Lys Val Ser Ile Gln Asp Val Asn  
 405 410 415

tcc gct atg gaa aat atc ttt caa aac aat ata aga tta acc ggg cat 1296  
 Ser Ala Met Glu Asn Ile Phe Gln Asn Asn Ile Arg Leu Thr Gly His  
 420 425 430

tta tta cct aat gga gaa 1314  
 Leu Leu Pro Asn Gly Glu  
 435

&lt;210&gt; 5

&lt;211&gt; 438

&lt;212&gt; PRT

&lt;213&gt; Ehrlichia canis

&lt;400&gt; 5

```

Met Met Lys Phe Phe Thr Cys Phe Phe Ile Val Phe Leu Thr Ile Ala
 1           5           10           15
Asn His Ala Leu Ser Phe Asn Ile Lys Val Thr His Glu Lys Leu Asp
          20           25           30
Asn Gly Met Glu Val Tyr Val Ile Pro Asn His Arg Ala Pro Ala Val
          35           40           45
Met His Met Val Leu Tyr Lys Val Gly Gly Thr Asp Asp Pro Val Gly
          50           55           60
Tyr Ser Gly Leu Ala His Phe Phe Glu His Leu Met Phe Ser Gly Thr
 65           70           75           80
Glu Lys Phe Pro Asn Leu Ile Ser Thr Leu Ser Asn Ile Gly Gly Asn
          85           90           95
Phe Asn Ala Ser Thr Ser Gln Phe Cys Thr Ile Tyr Tyr Glu Leu Ile
          100          105          110
Pro Lys Gln Tyr Leu Ser Leu Ala Met Asp Ile Glu Ser Asp Arg Met
          115          120          125
Gln Asn Phe Lys Val Thr Asp Lys Ala Leu Ile Arg Glu Gln Lys Val
          130          135          140
Val Leu Glu Glu Arg Lys Met Arg Val Glu Ser Gln Ala Lys Asn Ile
          145          150          155          160
Leu Glu Glu Glu Met Glu Asn Ala Phe Tyr Tyr Asn Gly Tyr Gly Arg
          165          170          175
Pro Val Val Gly Trp Glu His Glu Ile Ser Asn Tyr Asn Lys Glu Val
          180          185          190
Ala Glu Ala Phe His Lys Leu His Tyr Ser Pro Asn Asn Ala Ile Leu
          195          200          205
Ile Val Thr Gly Asp Ala Asp Pro Gln Glu Val Ile Thr Leu Ala Lys
          210          215          220
Gln Tyr Tyr Gly Lys Ile Pro Ser Asn Asn Lys Lys Pro Ser Ser Gln
          225          230          235          240
Val Arg Val Glu Pro Pro His Lys Thr Asn Met Thr Leu Thr Leu Lys
          245          250          255
Asp Ser Ser Val Glu Ile Pro Glu Leu Phe Leu Met Tyr Gln Ile Pro
          260          265          270

```

10

Asn Gly Ile Thr Asn Lys Asn Tyr Ile Leu Asn Met Met Leu Ala Glu  
           275                                  280                                  285  
 Ile Leu Gly Ser Gly Lys Phe Ser Leu Leu Tyr Asn Asp Leu Val Ile  
           290                                  295                                  300  
 Asn Asn Pro Ile Val Thr Ser Ile Lys Thr Asp Tyr Asn Tyr Leu Thr  
   305                                  310                                  315                                  320  
 Asp Ser Asp Asn Tyr Leu Ser Ile Glu Ala Ile Pro Lys Asn Gly Ile  
                                   325                                  330                                  335  
 Ser Thr Glu Ala Val Glu Gln Glu Ile His Lys Cys Ile Asn Asn Tyr  
                                   340                                  345                                  350  
 Leu Glu Asn Gly Ile Ser Ala Glu Tyr Leu Glu Ser Ala Lys Tyr Lys  
                                   355                                  360                                  365  
 Val Lys Ala His Leu Thr Tyr Ala Phe Asp Gly Leu Thr Phe Ile Ser  
                                   370                                  375                                  380  
 Tyr Phe Tyr Gly Met His Leu Ile Leu Gly Val Pro Leu Ser Glu Ile  
   385                                  390                                  395                                  400  
 Ser Asn Ile Tyr Asp Thr Ile Asp Lys Val Ser Ile Gln Asp Val Asn  
                                   405                                  410                                  415  
 Ser Ala Met Glu Asn Ile Phe Gln Asn Asn Ile Arg Leu Thr Gly His  
                                   420                                  425                                  430  
 Leu Leu Pro Asn Gly Glu  
                                   435

&lt;210&gt; 6

&lt;211&gt; 1407

&lt;212&gt; DNA

&lt;213&gt; Ehrlichia canis

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1407)

<223> Protein translated from 2,258 through 3,664  
 (ProB).

&lt;400&gt; 6

atg aga aac ata ttg tgt tac aca tta ata ttg att ttc ttt tca ttc 48  
 Met Arg Asn Ile Leu Cys Tyr Thr Leu Ile Leu Ile Phe Phe Ser Phe  
   1                                  5                                  10                                  15

aat aca tat gca aat gat ctc aat att aac ata aaa gaa gct aca act	96
Asn Thr Tyr Ala Asn Asp Leu Asn Ile Asn Ile Lys Glu Ala Thr Thr	
20 25 30	
aaa aat aaa ata cac tat cta tat gtt gaa cat cat aac cta cca aca	144
Lys Asn Lys Ile His Tyr Leu Tyr Val Glu His His Asn Leu Pro Thr	
35 40 45	
att tcc tta aaa ttt gca ttc aag aaa gca gga tac gct tat gat gcc	192
Ile Ser Leu Lys Phe Ala Phe Lys Lys Ala Gly Tyr Ala Tyr Asp Ala	
50 55 60	
ttt gat aag caa gga ctt gca tac ttt aca tca aaa ata tta aac gaa	240
Phe Asp Lys Gln Gly Leu Ala Tyr Phe Thr Ser Lys Ile Leu Asn Glu	
65 70 75 80	
gga tca aaa aac aac tat gct ctc agt ttt gca caa caa tta gaa ggc	288
Gly Ser Lys Asn Asn Tyr Ala Leu Ser Phe Ala Gln Gln Leu Glu Gly	
85 90 95	
aaa ggt ata gac tta aaa ttt gat ata gac cta gac aat ttt tat ata	336
Lys Gly Ile Asp Leu Lys Phe Asp Ile Asp Leu Asp Asn Phe Tyr Ile	
100 105 110	
tca tta aaa acc tta tca gaa aac ttt gaa gaa gcc cta gtt tta ctc	384
Ser Leu Lys Thr Leu Ser Glu Asn Phe Glu Glu Ala Leu Val Leu Leu	
115 120 125	
agt gat tgc ata ttc aac acc gtc aca gat caa gaa ata ttc aat aga	432
Ser Asp Cys Ile Phe Asn Thr Val Thr Asp Gln Glu Ile Phe Asn Arg	
130 135 140	
ata ata gca gaa cag att gca cat gtt aaa tca tta tat tct gct cct	480
Ile Ile Ala Glu Gln Ile Ala His Val Lys Ser Leu Tyr Ser Ala Pro	
145 150 155 160	
gaa ttt ata gct aca aca gaa atg aat cac gct ata ttc aaa ggg cac	528
Glu Phe Ile Ala Thr Thr Glu Met Asn His Ala Ile Phe Lys Gly His	
165 170 175	
cca tat tct aac aaa gtt tac ggg aca tta aat aca atc aat aat atc	576
Pro Tyr Ser Asn Lys Val Tyr Gly Thr Leu Asn Thr Ile Asn Asn Ile	
180 185 190	
aac cag gaa gac gtt gca tta tat ata aaa aat agt ttt gac aag gaa	624
Asn Gln Glu Asp Val Ala Leu Tyr Ile Lys Asn Ser Phe Asp Lys Glu	
195 200 205	
caa atc gtt atc agc gca gca gga gat gta gat cca aca cag cta tca	672
Gln Ile Val Ile Ser Ala Ala Gly Asp Val Asp Pro Thr Gln Leu Ser	
210 215 220	
aat tta cta gat aaa tat att ctt tcc aaa ttg cca tct ggt aat aac	720
Asn Leu Leu Asp Lys Tyr Ile Leu Ser Lys Leu Pro Ser Gly Asn Asn	
225 230 235 240	

aaa aat acc ata cca gat acg act gtt aat aga gaa gac aca tta tta	768
Lys Asn Thr Ile Pro Asp Thr Thr Val Asn Arg Glu Asp Thr Leu Leu	
245 250 255	
tat gta cag aga gat gta cca caa agt gtc ata atg ttt gct aca gac	816
Tyr Val Gln Arg Asp Val Pro Gln Ser Val Ile Met Phe Ala Thr Asp	
260 265 270	
aca gta cca tat cac agc aaa gac tat cat gca tca aac ttg ttc aat	864
Thr Val Pro Tyr His Ser Lys Asp Tyr His Ala Ser Asn Leu Phe Asn	
275 280 285	
act atg cta ggc gga tta agt ctc aat tca ata tta atg ata gaa tta	912
Thr Met Leu Gly Gly Leu Ser Leu Asn Ser Ile Leu Met Ile Glu Leu	
290 295 300	
aga gac aag tta gga tta aca tac cat agt agc agt tca cta tct aac	960
Arg Asp Lys Leu Gly Leu Thr Tyr His Ser Ser Ser Ser Leu Ser Asn	
305 310 315 320	
atg aat cat agt aat gtg cta ttt ggt aca ata ttc act gat aat acc	1008
Met Asn His Ser Asn Val Leu Phe Gly Thr Ile Phe Thr Asp Asn Thr	
325 330 335	
aca gta aca aaa tgt ata tcc gtc tta aca gat att ata gag cac att	1056
Thr Val Thr Lys Cys Ile Ser Val Leu Thr Asp Ile Ile Glu His Ile	
340 345 350	
aaa aag tat gga gtt gat gaa gac act ttt gca att gca aaa tct agt	1104
Lys Lys Tyr Gly Val Asp Glu Asp Thr Phe Ala Ile Ala Lys Ser Ser	
355 360 365	
att acc aac tct ttt att tta tct atg tta aat aac aat aat gtt agt	1152
Ile Thr Asn Ser Phe Ile Leu Ser Met Leu Asn Asn Asn Asn Val Ser	
370 375 380	
gag ata ttg tta agc tta caa tta cac gat cta gat ccg agt tat att	1200
Glu Ile Leu Leu Ser Leu Gln Leu His Asp Leu Asp Pro Ser Tyr Ile	
385 390 395 400	
aat aaa tac aat tct tac tac aaa gca ata aca ata gaa gaa gta aat	1248
Asn Lys Tyr Asn Ser Tyr Tyr Lys Ala Ile Thr Ile Glu Glu Val Asn	
405 410 415	
aaa att gcc aag aaa att tta tct aat gaa tta gta ata att gaa gta	1296
Lys Ile Ala Lys Lys Ile Leu Ser Asn Glu Leu Val Ile Ile Glu Val	
420 425 430	
gga aaa aac aat aac ata aat ggc aaa caa ata gat gct aaa aaa cac	1344
Gly Lys Asn Asn Asn Ile Asn Gly Lys Gln Ile Asp Ala Lys Lys His	
435 440 445	
ata cct tgg tta agt ata cag gtt att gta ttt act aca agt att cta	1392
Ile Pro Trp Leu Ser Ile Gln Val Ile Val Phe Thr Thr Ser Ile Leu	
450 455 460	

tta ggt tgt att aag  
 Leu Gly Cys Ile Lys  
 465

1407

&lt;210&gt; 7

&lt;211&gt; 469

&lt;212&gt; PRT

&lt;213&gt; Ehrlichia canis

&lt;400&gt; 7

Met	Arg	Asn	Ile	Leu	Cys	Tyr	Thr	Leu	Ile	Leu	Ile	Phe	Phe	Ser	Phe		
1				5					10					15			
Asn	Thr	Tyr	Ala	Asn	Asp	Leu	Asn	Ile	Asn	Ile	Lys	Glu	Ala	Thr	Thr		
			20					25					30				
Lys	Asn	Lys	Ile	His	Tyr	Leu	Tyr	Val	Glu	His	His	Asn	Leu	Pro	Thr		
		35					40					45					
Ile	Ser	Leu	Lys	Phe	Ala	Phe	Lys	Lys	Ala	Gly	Tyr	Ala	Tyr	Asp	Ala		
	50					55					60						
Phe	Asp	Lys	Gln	Gly	Leu	Ala	Tyr	Phe	Thr	Ser	Lys	Ile	Leu	Asn	Glu		
65					70					75					80		
Gly	Ser	Lys	Asn	Asn	Tyr	Ala	Leu	Ser	Phe	Ala	Gln	Gln	Leu	Glu	Gly		
			85						90					95			
Lys	Gly	Ile	Asp	Leu	Lys	Phe	Asp	Ile	Asp	Leu	Asp	Asn	Phe	Tyr	Ile		
			100					105					110				
Ser	Leu	Lys	Thr	Leu	Ser	Glu	Asn	Phe	Glu	Glu	Ala	Leu	Val	Leu	Leu		
		115					120					125					
Ser	Asp	Cys	Ile	Phe	Asn	Thr	Val	Thr	Asp	Gln	Glu	Ile	Phe	Asn	Arg		
		130				135					140						
Ile	Ile	Ala	Glu	Gln	Ile	Ala	His	Val	Lys	Ser	Leu	Tyr	Ser	Ala	Pro		
145					150					155					160		
Glu	Phe	Ile	Ala	Thr	Thr	Glu	Met	Asn	His	Ala	Ile	Phe	Lys	Gly	His		
				165				170						175			
Pro	Tyr	Ser	Asn	Lys	Val	Tyr	Gly	Thr	Leu	Asn	Thr	Ile	Asn	Asn	Ile		
			180					185					190				
Asn	Gln	Glu	Asp	Val	Ala	Leu	Tyr	Ile	Lys	Asn	Ser	Phe	Asp	Lys	Glu		
		195					200					205					
Gln	Ile	Val	Ile	Ser	Ala	Ala	Gly	Asp	Val	Asp	Pro	Thr	Gln	Leu	Ser		
		210				215					220						



14

Asn Leu Leu Asp Lys Tyr Ile Leu Ser Lys Leu Pro Ser Gly Asn Asn  
 225 230 235 240  
 Lys Asn Thr Ile Pro Asp Thr Thr Val Asn Arg Glu Asp Thr Leu Leu  
 245 250 255  
 Tyr Val Gln Arg Asp Val Pro Gln Ser Val Ile Met Phe Ala Thr Asp  
 260 265 270  
 Thr Val Pro Tyr His Ser Lys Asp Tyr His Ala Ser Asn Leu Phe Asn  
 275 280 285  
 Thr Met Leu Gly Gly Leu Ser Leu Asn Ser Ile Leu Met Ile Glu Leu  
 290 295 300  
 Arg Asp Lys Leu Gly Leu Thr Tyr His Ser Ser Ser Ser Leu Ser Asn  
 305 310 315 320  
 Met Asn His Ser Asn Val Leu Phe Gly Thr Ile Phe Thr Asp Asn Thr  
 325 330 335  
 Thr Val Thr Lys Cys Ile Ser Val Leu Thr Asp Ile Ile Glu His Ile  
 340 345 350  
 Lys Lys Tyr Gly Val Asp Glu Asp Thr Phe Ala Ile Ala Lys Ser Ser  
 355 360 365  
 Ile Thr Asn Ser Phe Ile Leu Ser Met Leu Asn Asn Asn Asn Val Ser  
 370 375 380  
 Glu Ile Leu Leu Ser Leu Gln Leu His Asp Leu Asp Pro Ser Tyr Ile  
 385 390 395 400  
 Asn Lys Tyr Asn Ser Tyr Tyr Lys Ala Ile Thr Ile Glu Glu Val Asn  
 405 410 415  
 Lys Ile Ala Lys Lys Ile Leu Ser Asn Glu Leu Val Ile Ile Glu Val  
 420 425 430  
 Gly Lys Asn Asn Asn Ile Asn Gly Lys Gln Ile Asp Ala Lys Lys His  
 435 440 445  
 Ile Pro Trp Leu Ser Ile Gln Val Ile Val Phe Thr Thr Ser Ile Leu  
 450 455 460  
 Leu Gly Cys Ile Lys  
 465

&lt;210&gt; 8

&lt;211&gt; 675

&lt;212&gt; DNA

&lt;213&gt; Ehrlichia canis

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(675)

&lt;223&gt; Protein translated from nucleotides 4,121 through 4,795 (ORF of unknown function).

&lt;400&gt; 8

atg gta tta ttt atg aaa gct cat agc aca agt ata cgg aac ttt cag	48
Met Val Leu Phe Met Lys Ala His Ser Thr Ser Ile Arg Asn Phe Gln	
1 5 10 15	
cct tta gaa aga gct gct ata atc att gca gtg tta ggt tta gct gca	96
Pro Leu Glu Arg Ala Ala Ile Ile Ile Ala Val Leu Gly Leu Ala Ala	
20 25 30	
ttc ttg ttt gct gct gct gcc tgc agt gat cgt ttc caa aga ttg caa	144
Phe Leu Phe Ala Ala Ala Ala Cys Ser Asp Arg Phe Gln Arg Leu Gln	
35 40 45	
tta aca aat cca ttt gta ata gca gga atg gtt ggc ctt gca gtt ctt	192
Leu Thr Asn Pro Phe Val Ile Ala Gly Met Val Gly Leu Ala Val Leu	
50 55 60	
tta gtt gct tcc tta aca gca gca tta agt ata tgc tta act aaa agt	240
Leu Val Ala Ser Leu Thr Ala Ala Leu Ser Ile Cys Leu Thr Lys Ser	
65 70 75 80	
aag caa gtc aca caa cat gct att aga cat cgc ttt gga tac gag tca	288
Lys Gln Val Thr Gln His Ala Ile Arg His Arg Phe Gly Tyr Glu Ser	
85 90 95	
agc act tct tct tct gta ctg ctt gca ata tca ata att tct tta tta	336
Ser Thr Ser Ser Ser Val Leu Leu Ala Ile Ser Ile Ile Ser Leu Leu	
100 105 110	
ctt gct gca gca ttt tgt gga aag ata atg ggt aat gac aac cca gat	384
Leu Ala Ala Ala Phe Cys Gly Lys Ile Met Gly Asn Asp Asn Pro Asp	
115 120 125	
cta ttc ttt agc aag atg caa gaa ctc tcc aat cca ctt gtt gtt gca	432
Leu Phe Phe Ser Lys Met Gln Glu Leu Ser Asn Pro Leu Val Val Ala	
130 135 140	
gct att gta gcc gtt tct gtt ttc cta ctc tca ttc gta atg tat gct	480
Ala Ile Val Ala Val Ser Val Phe Leu Leu Ser Phe Val Met Tyr Ala	
145 150 155 160	
gca aag aac att ata agt cca gat aaa caa act cac gtt att ata tta	528
Ala Lys Asn Ile Ile Ser Pro Asp Lys Gln Thr His Val Ile Ile Leu	
165 170 175	

16

tct aat caa caa act ata gaa gaa gca aaa gta gat caa gga atg aat 576  
 Ser Asn Gln Gln Thr Ile Glu Glu Ala Lys Val Asp Gln Gly Met Asn  
                   180                  185                  190

att ttg tca gca gta ctc cca gca gct ggc att gac atc atg act ata 624  
 Ile Leu Ser Ala Val Leu Pro Ala Ala Gly Ile Asp Ile Met Thr Ile  
                   195                  200                  205

gct tct tgt gac att tta gca gtg agc agc cgg gga tcc tct cag cat 672  
 Ala Ser Cys Asp Ile Leu Ala Val Ser Ser Arg Gly Ser Ser Gln His  
                   210                  215                  220

caa 675  
 Gln  
 225

&lt;210&gt; 9

&lt;211&gt; 225

&lt;212&gt; PRT

&lt;213&gt; Ehrlichia canis

&lt;400&gt; 9

Met Val Leu Phe Met Lys Ala His Ser Thr Ser Ile Arg Asn Phe Gln  
   1                  5                  10                  15

Pro Leu Glu Arg Ala Ala Ile Ile Ile Ala Val Leu Gly Leu Ala Ala  
                   20                  25                  30

Phe Leu Phe Ala Ala Ala Ala Cys Ser Asp Arg Phe Gln Arg Leu Gln  
                   35                  40                  45

Leu Thr Asn Pro Phe Val Ile Ala Gly Met Val Gly Leu Ala Val Leu  
                   50                  55                  60

Leu Val Ala Ser Leu Thr Ala Ala Leu Ser Ile Cys Leu Thr Lys Ser  
                   65                  70                  75                  80

Lys Gln Val Thr Gln His Ala Ile Arg His Arg Phe Gly Tyr Glu Ser  
                   85                  90                  95

Ser Thr Ser Ser Ser Val Leu Leu Ala Ile Ser Ile Ile Ser Leu Leu  
                   100                  105                  110

Leu Ala Ala Ala Phe Cys Gly Lys Ile Met Gly Asn Asp Asn Pro Asp  
                   115                  120                  125

Leu Phe Phe Ser Lys Met Gln Glu Leu Ser Asn Pro Leu Val Val Ala  
                   130                  135                  140

Ala Ile Val Ala Val Ser Val Phe Leu Leu Ser Phe Val Met Tyr Ala  
                   145                  150                  155                  160

gat	cag	gta	agt	aaa	tgg	tat	gta	gta	aat	ttg	ata	gga	gat	aaa	ggt	48
Asp	Gln	Val	Ser	Lys	Trp	Tyr	Val	Val	Asn	Leu	Ile	Gly	Asp	Lys	Gly	
1				5					10					15		
gta	ata	gag	ata	tta	agc	ttc	ttg	cgc	ttt	act	aca	gtg	tgg	aat	gct	96
Val	Ile	Glu	Ile	Leu	Ser	Phe	Leu	Arg	Phe	Thr	Thr	Val	Trp	Asn	Ala	
			20					25					30			
gga	att	agt	ttt	ggt	ata	tta	aat	aac	ttt	gaa	tat	agt	aat	gtt	gtt	144
Gly	Ile	Ser	Phe	Gly	Ile	Leu	Asn	Asn	Phe	Glu	Tyr	Ser	Asn	Val	Val	
		35					40					45				
ttt	tgt	agt	atc	tcg	att	ttg	att	act	tgt	ggt	tta	tgc	tac	tta	ttt	192
Phe	Cys	Ser	Ile	Ser	Ile	Leu	Ile	Thr	Cys	Val	Leu	Cys	Tyr	Leu	Phe	
	50					55					60					
ata	gta	cag	cca	cat	tat	aga	tta	cct	ctt	gta	atc	att	att	ggg	ggg	240
Ile	Val	Gln	Pro	His	Tyr	Arg	Leu	Pro	Leu	Val	Ile	Ile	Ile	Gly	Gly	
65					70					75					80	

18

tca ata gga aat atc ata gat aga ata aga tat ggt gct gtc tat gat 288  
 Ser Ile Gly Asn Ile Ile Asp Arg Ile Arg Tyr Gly Ala Val Tyr Asp  
                     85                    90                    95

ttt ata gat ttt tat atc aat aac tta cat tgg cct gta ttc aac ctg 336  
 Phe Ile Asp Phe Tyr Ile Asn Asn Leu His Trp Pro Val Phe Asn Leu  
                     100                    105                    110

gcg gat tct ttt ata ttt tta ggt ata gta ata ata atg gca aag agt 384  
 Ala Asp Ser Phe Ile Phe Leu Gly Ile Val Ile Ile Met Ala Lys Ser  
                     115                    120                    125

aat aac cac atg aaa caa att aac tgt aac tcc 417  
 Asn Asn His Met Lys Gln Ile Asn Cys Asn Ser  
                     130                    135

&lt;210&gt; 11

&lt;211&gt; 139

&lt;212&gt; PRT

&lt;213&gt; Ehrlichia canis

&lt;400&gt; 11

Asp Gln Val Ser Lys Trp Tyr Val Val Asn Leu Ile Gly Asp Lys Gly  
   1                    5                    10                    15

Val Ile Glu Ile Leu Ser Phe Leu Arg Phe Thr Thr Val Trp Asn Ala  
                     20                    25                    30

Gly Ile Ser Phe Gly Ile Leu Asn Asn Phe Glu Tyr Ser Asn Val Val  
                     35                    40                    45

Phe Cys Ser Ile Ser Ile Leu Ile Thr Cys Val Leu Cys Tyr Leu Phe  
                     50                    55                    60

Ile Val Gln Pro His Tyr Arg Leu Pro Leu Val Ile Ile Ile Gly Gly  
                     65                    70                    75                    80

Ser Ile Gly Asn Ile Ile Asp Arg Ile Arg Tyr Gly Ala Val Tyr Asp  
                     85                    90                    95

Phe Ile Asp Phe Tyr Ile Asn Asn Leu His Trp Pro Val Phe Asn Leu  
                     100                    105                    110

Ala Asp Ser Phe Ile Phe Leu Gly Ile Val Ile Ile Met Ala Lys Ser  
                     115                    120                    125

Asn Asn His Met Lys Gln Ile Asn Cys Asn Ser  
                     130                    135